

**METHOD FOR ASSISTING AN AUTOMATED VIDEO  
TRACKING SYSTEM IN REAQUIRING A TARGET**

**BACKGROUND OF THE INVENTION**

**1. Field of the Invention**

5           The present invention relates generally to video tracking systems and, more particularly, to a method for reacquiring a target in an automated video tracking system.

**2. Prior Art**

10           For purposes of this disclosure, automated tracking is defined as the automatic control of the Pan, Tilt and Zoom (PTZ) motors of a movable PTZ camera so as to keep the camera view centered on a designated, moving target. Automated tracking as defined is used in a  
15           number of different applications areas, such as surveillance and security monitoring. In this area, the target is usually a human.

          Automated tracking systems typically have several parts, target selection, model generation, and  
20           camera control. A target needs to be selected for tracking. This can be via an operator or via an automated motion detection module or other intruder detection system. An internal "model" of the appearance of the target is necessary to allow the tracking system  
25           to find the target in subsequent images. A camera motion

control model is necessary to determine how to move the camera to keep the target in the center of the field of view.

5 The present disclosure relates to the problem  
of target selection and more particularly, on reacquiring  
a target in an ambiguous situation where the automatic  
tracker loses the selected target. Identification of  
potential tracking candidates (i.e., a desired target) in  
a video scene is typically not part of the function of an  
10 automated tracking system. For instance, in the area of  
surveillance, target selection requires a lot of  
background knowledge about the objective of any  
surveillance application. What looks "suspicious" in one  
surveillance application, e.g. a retail store, may not  
15 look suspicious in another, e.g. a parking lot.

In some applications, any source of motion is  
suspicious, e.g., monitoring a warehouse at night. In  
that case, an intrusion detection sensor, or a motion  
sensor, could be used to designate a target for tracking.  
20 A more sophisticated automatic monitoring system, could  
be used to designate targets for certain other  
applications, as long as the rules to select targets can  
be clearly enumerated and implemented. However, in  
general, in most commercially available systems,  
25 especially in the domain of surveillance, it is expected  
that a human operator will indicate the target to the  
tracking system. Some commercial systems have been  
developed that fully automate the selection of the  
target, however, these systems are not robust enough to

handle all of the realistic and normal cases that may be encountered in all applications, particularly in surveillance. Furthermore, it is not always suitable to allow the tracking system to have full control of

5 selecting the target because it is possible that another moving target may become more interesting to track. Systems having automated target selection frequently run off of the target, mainly because of the uncontrollable environment (e.g., illumination conditions, multiple

10 people, etc.). Generally, the systems which employ automated target selection work better for applications where conditions are more predictable and less likely to change, such as video-conferencing, presentation, and learning and do not work well where conditions are less

15 predictable, such as surveillance.

In the systems which employ manual target selection, the operator selects the target by using a joystick to control the pan and tilt motors of a PTZ camera and possibly even the zoom motor of the PTZ

20 camera. The operator manipulating the joystick needs to be trained to correctly use the joystick because tracking in a three-dimensional environment can be very difficult, especially where the target does not have a predictable path and/or is moving rapidly.

25 When an operator designates a person in the video image as the tracking system's target, there is a subtle difference in meaning between the operator's and the tracking system's concept of the target. The operator is designating a person as the target, however,

the tracking system is simply accepting a region of the image as the target. Because of this, the operator may not be overly fussy about what part of the person he picks, since after all, its clear to any (human) observer  
5 which person he or she selects. Furthermore, the tracking system will form a target model based on exactly what image region the operator selected. As it has no independent knowledge of the desired target it cannot generalize beyond what it is told.

10 Therefore, there is a need in the art for method and apparatus which permits an operator of an automated video tracking system to take control of the same and reacquire a target when the video tracking system encounters a period of difficulty in tracking the  
15 target.

#### **SUMMARY OF THE INVENTION**

Therefore it is an object of the present invention to provide a method and apparatus for reacquiring a target in a video tracking system which  
20 resolves the problems with the prior art video tracking systems.

Accordingly, a method for reacquiring a target in an automated video tracking system is provided. The method comprises the steps of: (a) selecting a desired  
25 target to be tracked; (b) switching the automated video tracking system to an automatic mode to initiate a tracking sequence to automatically track the selected desired target; (c) switching the automated video

tracking system from an automatic mode to a manual mode if the automated video tracking system encounters a period of difficulty in tracking the desired target; (d) reacquiring the desired target in manual mode; and (e) 5 switching the automated video tracking system to the automatic mode for automatic tracking of the reacquired desired target without initiating a new tracking sequence.

Preferably, step (a) comprises centering the 10 desired target in a display of a scene including the desired target, step (b) comprises releasing control of an input device used to select the desired target, (c) comprises controlling an input device used to select the desired target, step (d) comprises centering the desired 15 target in a display of a scene including the desired target, and step (e) comprises releasing control of an input device used to reacquire the desired target.

Also provided is an apparatus for reacquiring a target in an automated video tracking system, where the 20 apparatus comprises: selecting means for selecting a desired target to be tracked; mode switching means for switching the automated video tracking system to and from one of an automatic mode to initiate a tracking sequence after target selection to automatically track the 25 selected desired target and a manual mode; reacquiring means for reacquiring the desired target in manual mode if the automated video tracking system encounters a period of difficulty in tracking the desired target; wherein after reacquiring the desired target the

automated video tracking system is switched back to automatic mode without initiating a new tracking sequence.

Preferably, the selecting means comprises an input device for centering the desired target in a display of a scene including the desired target, the mode selecting means comprises an input device where the automated video tracking system is switched to automatic mode by controlling an input device used to select the desired target and the automated video tracking system is switched to manual mode by releasing control of the input device, and the reacquiring means comprises an input device for centering the desired target in a display of a scene including the desired target. The apparatus preferably further comprises: a video camera for capturing video image data of a scene including the desired target; pan and tilt camera motors for controlling a pan and tilt, respectively of the video camera; and a video display for displaying the video image data; wherein the input device is a joystick operatively connected to the pan and tilt motors such that movement of the joystick controls the movement of the camera through the pan and tilt motors.

Still yet provided is an automated video tracking system for tracking and reacquiring a target. The automated video tracking system comprises: a video camera for capturing video image data of a scene including a desired target; pan and tilt camera motors for controlling a pan and tilt, respectively of the video

camera; a video display for displaying the video image data; selecting means for selecting the desired target to be tracked; mode switching means for switching the automated video tracking system to and from one of an  
5 automatic mode to initiate a tracking sequence after target selection to automatically track the selected desired target and a manual mode; reacquiring means for reacquiring the desired target in manual mode if the automated video tracking system encounters a period of  
10 difficulty in tracking the desired target; wherein after reacquiring the desired target the automated video tracking system is switched back to automatic mode without initiating a new tracking sequence.

Preferably, the selecting means comprises an  
15 input device for centering the desired target in the display, the mode selecting means comprises an input device where the automated video tracking system is switched to automatic mode by controlling an input device used to select the desired target and the automated video  
20 tracking system is switched to manual mode by releasing control of the input device, and the reacquiring means comprises an input device for centering the desired target in a display of a scene including the desired target. Preferably, the input device is a joystick  
25 operatively connected to the pan and tilt motors such that movement of the joystick controls the movement of the camera through the pan and tilt motors.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

These and other features, aspects, and advantages of the apparatus and methods of the present invention will become better understood with regard to  
5 the following description, appended claims, and accompanying drawings where:

Figure 1 illustrates a preferred video tracking system of the present invention.

Figure 2 illustrates a flowchart showing the  
10 steps of a preferred method for reacquiring a target in the video tracking system of Figure 1.

## **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT**

Although this invention is applicable to numerous and various types of video tracking systems, it  
15 has been found particularly useful in the environment of surveillance and security systems. Therefore, without limiting the applicability of the invention to surveillance and security systems, the invention will be described in such environment. Those skilled in the art  
20 will appreciate that the methods and apparatus of the present invention also have usefulness in such areas as videoconferencing and multi-modal interfaces for consumer devices.

Referring now to Figure 1, there is illustrated  
25 a preferred implementation of the video tracking system of the present invention, generally referred to by reference numeral 100. The apparatus 100 comprises a camera 102 for providing video image data of a scene 104



having a desired target 106 to be tracked. The camera 102 is preferably a PTZ camera having PTZ motors 108 for controlling the pan, tilt and zoom of the camera 102. Such camera and motors for their control are well known  
5 in the art.

The apparatus 100 further includes a display, such as a computer monitor 110 for displaying the video image data of the scene 104 from the camera 102, the monitor's display being referred to by reference numeral  
10 110a. An input device is used to select the desired target 106a in the video image data. Reference numeral 106 is used herein to indicate the actual target while reference numeral 106a indicates the image of the target as displayed on the monitor 110. The input device is  
15 preferably a joystick 112 connected to a computer processor 114 which also controls the PTZ motors 108 of the camera 102. However, any input device that is capable of selecting a target in the video image display can be utilized without departing from the scope or  
20 spirit of the present invention. Such other input devices can be a computer mouse, touchpad, touchscreen, touchpen, or even a keyboard 113 connected to the computer 114.

Once the operator has selected the target 106a,  
25 a tracking system 116 generates a model of the target 106a that can be used to locate the target 106a in successive frames of the video image data. Such tracking systems are well known in the art. Although shown separately in Figure 1, the tracking system 116 is

preferably implemented by software contained on a peripheral device (not shown) in the computer processor 114.

Typically, there are two *interconnected* ways in which the model is used, to distinguish the target 106a from the background scene or to distinguish the target 106a from other occluding targets. Because the model is gathered from the video image, it is clear that it can only contain information about appearance. This gives rise to the most important constraint limiting the behavior of automated tracking, referred to as the appearance constraint. In general, a target can only be successfully tracked if its appearance distinguishes it from other potential targets. In other words, if the target does not have something unique about its appearance within the kind of visual environments in which the tracker is operating, then it is not possible to build a unique "model" for that target.

Of course, a trained human observer is very good at picking up small clues from the visual image that are beyond the current state of the art in computer vision. For example, a surveillance operator can recognize a target from a partial view of his face, or by noting that a person has a unique way of holding his head. In general, an automated tracker is limited to looking at features of the video image such as whether or not a person has moved since the last frame, or the color composition of the clothes of an individual. An automated tracker is also typically limited by the

resolution of the camera image and by the lighting conditions in the field of view. Human operators can handle a wide range of different lighting conditions and variation of lighting within a scene. For this reason,  
5 even distinguishable targets may occasionally fail to be tracked.

Once a model of the target 106a is generated, a controller 118 is then instructed by the tracking system 116 to control the PTZ motors 108 to move the camera 102  
10 to keep the selected target 106 centered in the field of view of the camera 102. Such controllers 118 are also well known in the art. The controller 118, like the tracking system 116, is preferably implemented by software contained on a peripheral device on the computer  
15 114. Two general approaches for controlling the camera 102 that are widely used in the prior art include a discrete approach in which the camera 102 is moved from time to time to keep the target 106 centered and a continuous approach in which the camera 102 is moved to  
20 keep the target 106 continuously centered.

It may appear that the continuous case is simply the discrete case where the period between camera movements approaches zero. However, in general, both the tracking system 116 and the camera controller 118 needed,  
25 will be quite different for each of these. In the discrete approach, the tracking system 116 needs to have enough information about the camera 102 to accurately move it a large distance to re-center the target 106. Typically this means knowing the Pan, Tilt and Zoom

settings of the camera 102 at all times, and being able to command the camera 102 to go to specific settings. Advantages of this approach include the fact that the camera 102 is only moved infrequently, and when it does  
5 move it can move very quickly and hence keep up with a very fast target 106. Disadvantages include the fact that the target 106 is rarely centered, that the camera PTZ position needs to be known, and that the fast moves may disorient an operator.

10 In the continuous approach, the tracking system needs to move the camera 102 a small distance but at a high rate. Although position control could be used to do this, the resultant motion would not be smooth. Instead velocity control of the Pan, Tilt and Zoom settings of  
15 the camera 102 is well suited to performing smooth motion. No position feedback is necessary since velocities are recalculated so frequently. However, it is necessary that the image processing component of tracking be capable of dealing with images that are taken during  
20 camera motion (since the camera is almost always in motion). Advantages of this approach include the fact that the target 106 is always well centered. Disadvantages include the fact that the camera 102 is always moving if the target 106 is moving and that a fast  
25 target 106 may be lost.

Whichever approach is used to control the camera 102, the results of which are fed to the controller 118 which determines the amount of camera movement necessary to keep the target 106 centered in the

field of view of the camera 102. The controller 118 outputs an appropriate signal to the PTZ motors 108 of the camera 102 to carry out its instructions to keep the target 106 centered in the camera's 102 field of view.

5 Such systems are well known in the art.

A further aspect of the methods and apparatus of the present invention will now be described with reference to the apparatus of Figure 1 and the flowchart of Figure 2. Figure 2 illustrates a method for  
10 reacquiring a target 106 in an automated video tracking system, the method generally referred to by reference numeral 200.

At step 202, a desired target 106 to be tracked is selected. This can be done by any means known in the  
15 art. Preferably, the target 106 is selected by centering the target 106a in the monitor's display 110a by manipulating the joystick 112 to control at least the pan and tilt motors 108 of the camera 102, and possibly the zoom motor 108. At step 204, after selecting a desired  
20 target 106a, the automated video tracking system is switched to an automatic mode to initiate a tracking sequence to automatically track the selected desired target 106a. Switching to automatic mode can be achieved manually such as by selecting a button on a user  
25 interface (not shown) in the monitor's display 110a. However, it is preferred that the switch to automatic mode occurs automatically, preferably upon releasing control of the joystick 112 or other input device used to select the desired target 106a.

At step 206 it is determined if the video tracking system has encountered a period of difficulty and has lost the selected target 106a or may be in jeopardy of losing the selected target 106a. If not, the method proceeds along path 206a and the automated video tracking system continues to automatically track the selected target until the method is terminated, the selected target 106a leaves the scene 104, or it is no longer desired to track the selected target 106a, all of which are shown schematically as steps 214 and 216. If the operator perceives a difficulty has or is about to be encountered, the method proceeds along path 206b by switching from automatic mode to manual mode at step 208. As with the switching into automatic mode, switching from automatic mode to manual mode can be manually done by the operator or automatically upon taking control of the joystick 112 or other input device used to select the desired target 106a.

Once in manual mode, the operator reacquires the desired target at step 210. It is preferred that the target is reacquired in a simpler manner as compared to the way it is initially selected, namely, by centering the desired target 106a in the monitor's display 110a of the scene 104 by manipulating the joystick 112 or other input device. However, those skilled in the art will appreciate that the desired target can be selected and/or reacquired by any means known in the art without departing from the scope or spirit of the present invention.

Once the target is reacquired, the automated video tracking system is switched back to the automatic mode at step 212 where desired target is automatically tracked without initiating a new tracking sequence. It is important for the operator to quickly reacquire the desired target 106a so that a new tracking sequence is not initiated after the target 106a is reacquired and automatic tracking restarted. Preferably, automatic tracking restarts by either a manual instruction from the user, as discussed above, or preferably automatically by releasing control of the joystick 112 or other input device used to reacquire the desired target 106a. After switching back to automatic mode, the automatic video tracking system continues to track the reacquired target 106a until the target leaves the scene, tracking of the reacquired target is no longer desired, or another area of difficulty is encountered by the automatic tracking system, all of which are shown as steps 214 and 216.

As discussed above, when the target 106a is selected, a computer model is build to represent the appearance of that target 106a. During tracking of the target 106a, whenever the tracker finds a part of the image that matches to the target model, it preferably computes a number which represents how well the target 106a matches the model. This number can vary for example from 0% match to 100% match. Where the 100% indicates that the target matches the model completely. This value is called the confidence value. A control can also be provided to indicate a threshold value for the confidence. Thus, should the model match the target with

less than this threshold value, the operator is preferably warned by a signal of some form that the confidence is lower than the threshold and the tracking is about to fail. In such a situation, the operator may  
5 take over for the system for the time it takes for the target to pass the source of difficulty by reacquiring the target, after which control is given back to the automatic tracking system.

In summary, a means is used for switching an  
10 automatic video tracking system between two modes, automatic and manual, such as by controlling a joystick 112 or other input device. An operator initializes automatic tracking by centering the target in the image using the joystick 112 or other input device. After  
15 releasing the joystick 112, the automatic tracker enters automatic mode, locks on the selected target, and automatically tracks the selected target. If a particular situation happens which is difficult for the automatic video tracking system to resolve, for instance  
20 multiple people passing by the same location at the same time as the selected target, the operator may take over for the system for the time it takes for the target to pass the source of difficulty.

Those skilled in the art will appreciate that  
25 the apparatus and methods of the present invention allows an operator of a video tracking system, when necessary, to take control of the camera 102 by simply manipulating the joystick 112 or other input device. Thus, when something ambiguous to the video tracking system occurs



which is difficult for the automatic video tracking system to resolve, the operator can take control of the camera 102 back for a certain period of time to reacquire the target. This greatly simplifies the training needed  
5 by an operator and the operator's task in tracking the target.

While there has been shown and described what is considered to be preferred embodiments of the invention, it will, of course, be understood that various  
10 modifications and changes in form or detail could readily be made without departing from the spirit of the invention. It is therefore intended that the invention be not limited to the exact forms described and illustrated, but should be constructed to cover all  
15 modifications that may fall within the scope of the appended claims.